

2

AVANCES Y LIMITACIONES EN LA EVALUACIÓN DEL APRENDIZAJE A PARTIR DEL PROCESO DE CONVERGENCIA. VISIÓN DOCENTE Y DISCENTE EN LOS GRADOS DE EDUCACIÓN INFANTIL Y PRIMARIA

(ASSESSMENT PROGRESS AND LIMITATIONS AFTER THE CONVERGENCE PROCESS: STUDENT AND TEACHER PERCEPTION ON THE PRE-SERVICE TEACHER TRAINING)

Ainhoa Bilbao Martínez

Escuela Universitaria de Magisterio Begoñako Andra Mari, BAM (centro adscrito a la Universidad de Deusto)

Aurelio Villa Sánchez

Profesor Emérito en la Universidad de Deusto

DOI: 10.5944/educXX1.19976

Cómo referenciar este artículo/How to reference this article:

Bilbao Martínez, A. y Villa Sánchez, A. (2019). Avances y limitaciones en la evaluación del aprendizaje a partir del proceso de convergencia. Visión docente y discente en los grados de Educación Infantil y Primaria. *Educación XX1*, 22(1), 45-69, doi: 10.5944/educXX1.19976

Bilbao Martínez, A. & Villa Sánchez, A. (2019). Avances y limitaciones en la evaluación del aprendizaje a partir del proceso de convergencia. Visión docente y discente en los grados de Educación Infantil y Primaria. *Educación XX1*. [Assessment progress and limitations after the convergence process: student and teacher perception on the pre-service teacher training]. *Educación XX1*, 22(1), 45-69, doi: 10.5944/educXX1.19976

RESUMEN

Estudios recientes revelan avances puntuales en algunas disfuncionalidades clásicas de la evaluación del aprendizaje universitario. Su desarrollo se vincula a la construcción de un nuevo paradigma evaluativo, ligado al Espacio Europeo de Educación Superior y caracterizado en cinco retos de mejora: continuidad sensible de la evaluación, orientación estratégica al aprendizaje, tareas e instrumentos para el desarrollo y acreditación de competencias, explicitación e interiorización de criterios y condiciones evaluativas y participación discente. Este trabajo contrasta la visión de los mismos en una muestra de estudiantes y profesorado de los Grados de

Educación Infantil y Primaria. La indagación se ha realizado por medio de un cuestionario, basado en rúbricas que detallan distintos grados de aceptación y cumplimiento en cada uno de los retos. Los resultados revelan que los cinco se juzgan relevantes, sin que se produzcan discrepancias en la aceptación declarada por docentes y discentes. Las diferencias detectadas dentro del intervalo de confianza tienen un efecto de tamaño pequeño y hacen referencia al grado de cumplimiento percibido en tres retos en particular. Al tratar la Explicitación de la evaluación y su Continuidad sensible, el alumnado describe un cumplimiento significativamente inferior, asociado a una menor explotación de la funcionalidad pedagógica de la evaluación. Los estudiantes reconocen que la evaluación no solo se realiza al final, aunque tampoco creen que se desarrolle de manera sensible e integrada como afirma el profesorado. Asimismo, admiten disponer de información sobre las condiciones evaluativas, pero niegan la realización de tareas que contribuyan a su clarificación. Al valorar la participación discente en la evaluación, ambos sectores de la muestra reconocen una consecución insuficiente, aunque los promedios de respuesta del alumnado son significativamente menores. Estos resultados sugieren que a pesar de los avances advertidos en la forma, en el fondo la evaluación sigue representando una herramienta principalmente acreditativa.

PALABRAS CLAVE

Evaluación; Enseñanza Superior; Evaluación formativa; Evaluación continua; Autoevaluación; Evaluación por pares; Coevaluación.

ABSTRACT

Recent studies reveal punctual advances in some of the classical malfunctions of the university learning assessment. Its development is linked to the construction of a new assessment paradigm, based on the European Higher Education Area and featured by five main improvement challenges: Sensitive continuity of the evaluation, Strategic orientation to learning, Tasks and instruments for competency development and accreditation, Transparency and deep understanding of the assessment criteria and Student participation. This paper contrasts the perception of these challenges in a sample of students and teachers involved in educational degrees. The inquiry was carried out by means of a questionnaire, based on rubrics which detail five acceptance and fulfilment levels in relation to each challenge. The results disclose that the five of them are considered important and there is no divergence in the acceptance declared by both teachers and students. The differences found in the confidence interval have a small effect and refer to the degree of compliance particularly noticed in three of the challenges. Students describe a significantly inferior fulfilment when addressing the transparency in assessment criteria and the sensitive continuity, connected to the lower use of the formative function of the assessment. Students recognise

that the assessment does not only come at the end, but they do not think it is developed in a sensitive and integrated way, as teaching staff state. They also acknowledge having information about the assessment criteria, but they reject the carrying out of tasks focused on explaining the same. Both sectors of the sample recognise insufficient achievement in relation to the student participation, although the average from the students' answers is significantly inferior. These results suggest that despite the advances observed on its form, the assessment still represents a mainly certifying tool.

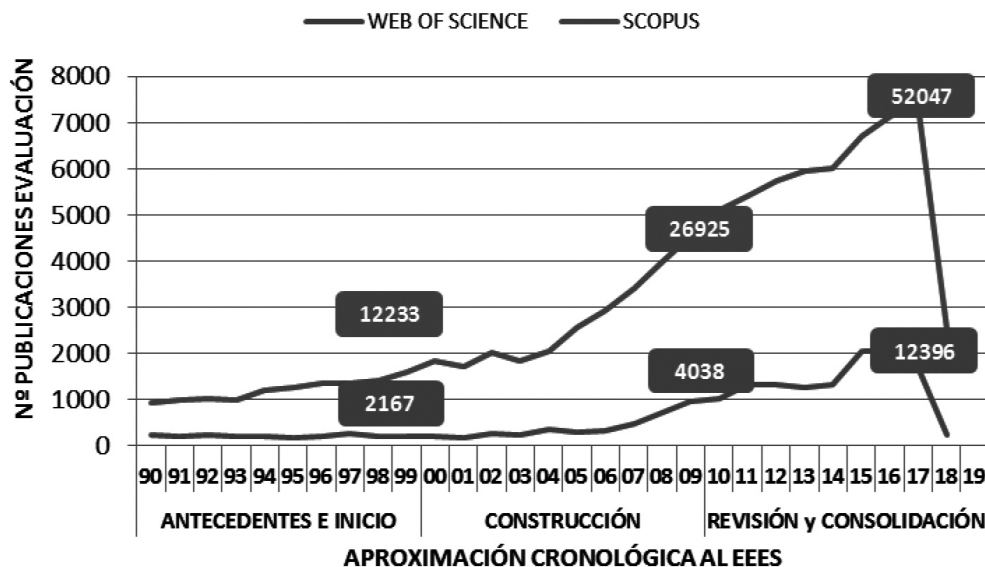
KEY WORDS

Assessment; Graduate Study; Formative evaluation; Continuous assessment; Self-assessment; Peer-assessment; Co-assessment.

INTRODUCCIÓN

La evaluación del aprendizaje universitario representa un foco clásico de atención dentro de la investigación educativa. Lejos de agotar su interés, la búsqueda de modelos didácticos acordes al paradigma formativo del Espacio Europeo de Educación Superior (EEES) ha desembocado en una nueva fase de efervescencia. Esta se hace patente en las dos últimas décadas, donde se aprecia un aumento reseñable del número de publicaciones centradas en la temática evaluativa (ver *Figura 1*).

Además de un interés creciente, esta evolución sugiere un cierto avance en la forma de realizar la evaluación. Así se constata al comparar las prácticas evaluativas descritas en el contexto universitario español antes de iniciarse el proceso de convergencia (1990-1999), en su construcción (2000-2009) o en la fase actual de revisión y consolidación (2010-2019).



Nota. La búsqueda se centra en dos bases de datos e incluye aquellos trabajos que lleven en el título el término Evaluation o Assessment. En la Web of Science se ha limitado la indagación a la categoría Education, educational research; en Scopus al área de las Ciencias Sociales. Consulta realizada el 24/04/2018.

Figura 1. Publicaciones sobre evaluación del aprendizaje en los antecedentes (1990-1999), la construcción (2000-2009) y la revisión y consolidación del EEES

El punto de partida: la evaluación del aprendizaje en la universidad española antes del EEES

Los diagnósticos efectuados en la fase previa del proceso de convergencia devuelven una imagen de la evaluación del aprendizaje negativa, repleta de síntomas de disfuncionalidad. Centrados en el punto de vista del alumnado (Fernández, 1989; Flórez, 1999), del profesorado (Rodríguez, 1998; Trillo y Porto, 1999) o en el contraste de ambos (Tejedor, 1999), sus conclusiones convergen en la descripción de una evaluación autocrática, final y acreditativa, poco explicitada y de escaso impacto formativo. Estas limitaciones se encuentran ampliamente referidas en la literatura específica, calificadas como patologías o esquizofrenias evaluativas (López, 2004). Fuera del ámbito nacional, también se produce gran concordancia al enumerar las principales limitaciones en las que se desarrolla la evaluación del aprendizaje universitario (Stiggins, 1999).

La construcción del EEES y la promesa de un nuevo paradigma evaluativo

Avanzar en un nuevo paradigma formativo es incompatible con la persistencia de los modelos de evaluación tradicionales (Cano y Ion, 2012). Desde esa perspectiva, la apuesta por una formación universitaria centrada en el aprendizaje de competencias exige el desarrollo de nuevas prácticas evaluativas que la sustenten. Su revisión debe transitar sobre cinco grandes líneas de avance y mejora, derivadas de las demandas formativas del EEES:

- La exploración de nuevas técnicas e instrumentos que aporten evidencias sobre el desarrollo de competencias, sin limitarse al dominio de contenidos aislados.
- La explicitación de los aprendizajes a realizar y de las condiciones para su evaluación, de modo que se avance en prácticas evaluativas más transparentes, que contribuyan a la clarificación e interiorización de los indicadores de logro definidos.
- La ampliación de los agentes intervinientes en la evaluación, de modo que el alumnado ejerza activamente el protagonismo que le corresponde en la formación y disponga de recursos para auto-regular su aprendizaje.
- La explotación de la funcionalidad pedagógica de la evaluación, de modo que sirva de herramienta didáctica para el desarrollo y fortalecimiento del aprendizaje, no solo para su acreditación.
- Unido a ello, el desarrollo de sistemas de evaluación continua, que proporcionen evidencias sobre el desarrollo de competencias en las distintas fases del proceso formativo, sin ejercer una presión excesiva que perjudique el aprendizaje.

Ampliamente representados en la literatura específica, los retos de mejora enumerados conectan con la superación de algunas de las disfuncionalidades más clásicas atribuidas a la evaluación. Aunque su formulación parta del proceso de convergencia, las orientaciones evaluativas resultantes guardan gran correspondencia con las definidas desde la perspectiva teórica de la evaluación *para* el aprendizaje (Palacios y López, 2013; Pereira, Flores, y Niklasson, 2016). Coincidiendo con su difusión, los estudios referidos a esta fase de construcción comienzan a mostrar ciertos cambios en la praxis evaluativa:

- Se reduce el predominio de la evaluación final, reconocido por buena parte del profesorado en los primeros años de la década (Nieto, 2000; Pérez, Reyes, Palma, y Rafel, 2000). Los programas docentes y la normativa que regula su ejercicio en la universidad comienzan

- a definir la evaluación como un proceso continuo, presente en las distintas fases del ciclo formativo (Ibarra y Rodríguez, 2010; Padilla, Rodríguez, Gil, López, y Clarés, 2010). Asimismo, aumentan los estudios que recogen el compromiso práctico del profesorado en la articulación de sistemas de evaluación continua (Arribas, Carabias, y Monreal, 2010; Olmos y Rodríguez, 2010).
- A medida que avanza la década, se detecta la proliferación de nuevas herramientas evaluativas, orientadas al desarrollo y acreditación de competencias (Buscà, Pintor, Martínez, y Peire, 2010; Ibarra y Rodríguez, 2010). Sin embargo, persiste el protagonismo del examen, típicamente centrado en tareas de evocación de contenidos, más que en su aplicación y transferencia (Cid-Sabucedo, Pérez-Abellás, y Zabalza, 2009).
 - Frente a falta de información acusada en la fase anterior (Fernández, 1989; Rodríguez, 1998), las programaciones docentes y la normativa evaluativa muestran un avance en la explicitación y transparencia de los criterios y condiciones que rigen la evaluación del aprendizaje (Ibarra y Rodríguez, 2010; Padilla et al., 2010).
 - Aunque la heteroevaluación docente sigue representando la principal modalidad evaluativa, se acusa un cierto avance en el desarrollo de prácticas evaluativas de autoevaluación y coevaluación (Álvarez, 2008; Bretones, 2008; Brew, Riley, & Walta, 2009).

Revisión y consolidación: la frontera del 2020

Inmersos en la recta final de la fase de revisión y consolidación, la evaluación del aprendizaje permanece entre los grandes retos pedagógicos a superar. El número de publicaciones sobre el tema se sigue incrementando. Continúan apareciendo estudios que ahondan en los cambios iniciados en la fase anterior, especialmente en lo que concierne a la continuidad, la instrumentación, los agentes y la explicitación de la evaluación. Sin embargo, a pesar del avance constatado respecto a las dos últimas décadas, las áreas de mejora atribuidas a la evaluación siguen poniendo de manifiesto la necesidad de reforzar su funcionalidad pedagógica (Canabal y Margalef, 2017; Gutiérrez, Pérez, y Pérez, 2013; Hortigüela, Pérez, y Abella, 2015a; Villa, Campo, Arranz, Villa, y García, 2013). La aproximación al EEES parece haber propiciado cierta evolución pero, en esencia, la evaluación sigue presentándose como un elemento prioritariamente final y acreditativo, asociado usualmente a exámenes y sensaciones negativas.

- Consolidando la tendencia iniciada en la década de la anterior, estudios recientes señalan la generalización de los sistemas de

- evaluación continua (Fraile, López, Castejón, y Romero, 2013; Hortigüela, Pérez, y López, 2015b). Sin embargo, existen indicios que llevan a pensar que el incremento en la frecuencia evaluativa se ha desarrollado al margen de su misión formativa (Hamodi y López, 2012; Hortigüela et al., 2015a; Villa et al., 2013). Esta circunstancia suscita en la literatura una inquietud creciente, ante los riesgos de incurrir en sistemas que hagan de la evaluación continua una sucesión de demandas acreditativas, poco sensibles con el proceso de construcción del aprendizaje discente (Monereo y Castelló, 2009; Sabirón y Arraiz, 2013).
- Estudios desarrollados en esta última década revelan que estudiantes y profesorado admiten la diversidad metodológica en la práctica evaluativa, aunque no su aplicación normalizada (Gutiérrez, Pérez, Pérez, y Palacios, 2011; Martínez, Castejón, y Santos, 2012; Martínez-Mínguez y Flores, 2014; Pereira et al., 2016). Al igual que en la fase de construcción, el examen conserva su peso y presencia prioritaria en los sistemas evaluativos empleados (Gil-Flores, 2012). Le siguen los informes o trabajos escritos, que comparten la tendencia de centrar la evaluación en procesos de elaboración teórica, más que en capacidades cognitivas de orden superior (Asún, Romero, y Chivite, 2017; Christoforidou, Kyriakides, Antoniou, & Creemers, 2014).
 - La investigación específica pone de manifiesto discrepancias entre docentes y discentes a la hora de juzgar el grado de explicitación de las condiciones evaluativas (Gutiérrez et al., 2013; Martínez et al., 2012; Martínez-Mínguez y Flores, 2014). En términos generales se reconoce una correcta labor de difusión; incluso para el alumnado, que ostenta una visión más crítica de la realidad evaluativa, la claridad de la información destaca entre las mayores fortalezas (Gil-Flores, 2012). Sin embargo, el trabajo específico sobre los criterios de evaluación comunicados sigue sin figurar entre las prácticas habituales (Rodríguez, Ibarra, Gallego, López, y Quesada, 2012).
 - A pesar del desarrollo y proliferación experimentados en la década anterior, estudiantes y profesorado consideran que las fórmulas de evaluación participativa siguen representando una alternativa puntual y minoritaria (Gallardo-Fuentes, López-Pastor, y Carter, 2017; Gutiérrez et al., 2011; Rodríguez et al., 2012). Este predominio de la heteroevaluación docente se considera inconsistente con los planteamientos del EEES (Bretones, 2008; Ibarra y Rodríguez, 2010; Sayós, 2010) y tampoco contribuye a la autorregulación del aprendizaje (Brown, 2015; Hortigüela et al., 2015b).

En suma, los avances descritos señalan cambios puntuales en la práctica evaluativa, aunque en el fondo siguen siendo reconocibles los

modelos de evaluación tradicional y sus clásicas limitaciones (Asún et al., 2017; Palacios y López, 2013). A pocos años de rebasar la fase de Revisión y consolidación, esta conclusión debería conducir a la detección de nuevas líneas de actuación para que el paradigma evaluativo en la universidad realmente consolide el paradigma de aprendizaje que subyace al EEES.

MÉTODO

Objetivo, diseño y participantes

El trabajo presentado tiene por objeto contrastar el punto de vista que manifiestan alumnado y profesorado con respecto a la aceptación y el grado de cumplimiento de los principales retos evaluativos ligados al paradigma formativo del EEES (Continuidad sensible de la evaluación, Orientación estratégica al aprendizaje, Tareas e instrumentos para la evaluación de competencias, Explicitación e interiorización de criterios y Participación discente).

Su estudio se ha efectuado por medio de un diseño no experimental de finalidad descriptiva. Desarrollado en el marco institucional de la Universidad de Deusto (UD) y el centro adscrito Begoñako Andra Mari (BAM) de Bilbao, representa un estudio de caso focalizado en los Grados de Educación Infantil y Primaria. La población de referencia la conforman los estudiantes (N = 1693) y profesores (N = 67) adscritos a ambos centros y titulaciones.

Los datos presentados proceden de una muestra productora compuesta por 704 estudiantes y 51 docentes. Su selección se ha llevado a cabo por medio de un muestreo no probabilístico accidental, logrando una representación total del 76,5% en la población docente y del 41,5% en la del alumnado.

Variables e instrumento

Desde un enfoque metodológico cuantitativo, el análisis atiende a la visión que manifiestan estudiantes y profesorado sobre el grado de aceptación y el cumplimiento en los cinco retos evaluativos destacados.

- La aceptación de los retos se determina por medio de una escala ordinal, configurada en función del grado de importancia atribuida: 1 (*Irrelevante*), 2 (*Poco relevante*), 3 (*Relevancia moderada*), 4 (*Relevante*) y 5 (*Muy relevante*).

- Para cuestionar el grado de cumplimiento percibido, por cada reto se ofrece una rúbrica con descriptores específicos para cinco niveles de logro: 1 (*Inadecuado*), 2 (*Poco adecuado*), 3 (*Adecuado*), 4 (*Muy adecuado*) y 5 (*Excelente*). La Tabla 1 resume su definición operativa, que parte del cuestionario para el diagnóstico de la competencia evaluativa presentado por Bilbao (2015). Dicha herramienta responde a un modelo teórico que delimita ocho dimensiones clave en la praxis evaluativa, desarrolladas por medio de un sistema de 19 indicadores y sus respectivos descriptores de logro. Los cinco retos estudiados hacen referencia solo a tres de ellas (planificación estratégica, clarificación de criterios y condiciones, y evaluación participativa)¹.

Tabla 1
 Niveles de consecución en los retos evaluativos analizados. Descriptores de logro 1 (*Inadecuado*), 2 (*Poco adecuado*), 3 (*Adecuado*), 4 (*Muy adecuado*) y 5 (*Excelente*) (Bilbao, 2015)

Continuidad sensible. El sistema de evaluación recaba evidencias de aprendizaje en las distintas fases del ciclo formativo, de forma sensible a la construcción que efectúan las/os estudiantes (sin interferir o perjudicarlo).	
1. Final	Las tareas evaluativas se concentran básicamente en una determinada fase del proceso formativo (por lo general, en el tramo final).
2. Final y proceso	Las tareas evaluativas se concentran fundamentalmente en dos de las fases del proceso formativo (por lo general, en las fases central y final).
3. Continúa: no atomizada	Incluye tareas evaluativas al comienzo del proceso formativo, durante su desarrollo y al finalizarlo, pero sin diseccionar o atomizar el aprendizaje.
4. Continúa: intensidad graduada	Incluye tareas evaluativas en las distintas fases del ciclo formativo, determinando la concentración e intensidad de las mismas en función del proceso de aprendizaje del alumnado (momentos críticos, tiempos de asimilación, etc.).
5. Continúa: intensidad graduada y coordinada	Incluye tareas evaluativas en las distintas fases del ciclo formativo, determinando la concentración e intensidad de las mismas en función del proceso de aprendizaje del alumnado y en coordinación con otras asignaturas (previsión global de la curva de esfuerzo y de trabajo).

Orientación estratégica al aprendizaje. El sistema de evaluación combina estratégicamente sus funciones pedagógicas y acreditativas para orientar al alumnado en el logro de las competencias previstas en el currículo.	
1. Acreditativa	Prácticamente la totalidad de las actividades evaluativas realizadas cuentan para nota.
2. Predominio acreditativo	Combina actividades evaluativas para mejorar el aprendizaje con otras destinadas a determinar la calificación, pero no de forma sistemática ni proporcionada (la mayor parte cuenta para nota).
3. Combinación estratégica	Combina deliberadamente tareas de evaluación acreditativa y pedagógica para lograr el desarrollo de las competencias previstas, no solo su medición.
4. Regulación estratégica: tiempo y esfuerzo	La combinación deliberada de tareas pedagógicas y acreditativas de evaluación lleva a las/os estudiantes a dedicar el tiempo y esfuerzo que requiere su aprendizaje.
5. Regulación estratégica: gestión del tiempo y esfuerzo	La combinación deliberada de tareas pedagógicas y acreditativas de evaluación lleva al alumnado a dedicar el tiempo y esfuerzo que requiere su aprendizaje y a gestionarlo de manera efectiva a lo largo del proceso formativo.
Tareas e instrumentos para la evaluación de competencias. El sistema evaluativo dispone tareas e instrumentos adecuados al contexto formativo y pertinentes, para obtener evidencias sobre el aprendizaje de las competencias previstas en el currículo.	
1. Evaluación de contenidos	Emplea instrumentos y tareas que atienden exclusivamente el dominio de contenidos (reproducción, descripción teórica, aplicación descontextualizada), no el desarrollo de competencias (aplicación en situaciones reales o simuladas).
2. Aproximación competencias	Intercala instrumentos y tareas para evaluar las competencias, con otras centradas exclusivamente en el dominio de contenidos.
3. Evaluación competencias	Emplea instrumentos y tareas que exigen aplicación y transferencia para evaluar el grado de desarrollo de las competencias.
4. Evaluación diversa de competencias	Emplea instrumentos y tareas que, para evaluar el grado de desarrollo de las competencias, exigen aplicación y transferencia en diversas situaciones.
5. Evaluación contextualizada competencias	Emplea instrumentos y tareas que, para evaluar el grado de desarrollo de las competencias, exigen aplicación y transferencia en situaciones diferentes y próximas a la realidad profesional.

Explicitación e interiorización de criterios. El sistema de evaluación incorpora actividades específicas para que, durante el proceso formativo, el alumnado comprenda los aprendizajes a realizar y los criterios para su valoración.	
1. Explicitación inicial	Los criterios y condiciones que rigen el proceso de evaluación del aprendizaje se detallan por escrito al comienzo de curso en la guía docente y/o programa de la asignatura.
2. Explicitación y presentación	Al comienzo del curso se reserva un espacio para exponer y explicar los criterios y condiciones de evaluación detallados por escrito.
3. Explicitación, presentación y clarificación	Durante el curso se realizan actividades para clarificar el significado de los criterios de evaluación explicitados y facilitar la comprensión de los logros a realizar (recordatorios, ejemplos, aclaración de significados, etc.).
4. Explicitación, presentación y clarificación apoyada	Durante el curso se proporcionan al alumnado herramientas o materiales de apoyo que clarifican el significado de los criterios de evaluación explicitados y su transferencia a las tareas planteadas (rúbricas, listas de control, modelos, etc.).
5. Actividades específicas interiorización	Durante el curso se realizan actividades creadas expresamente para que las/os estudiantes apliquen las herramientas o materiales de apoyo proporcionados y profundicen en su comprensión hasta interiorizarlos.
Participación discente. El sistema de evaluación incorpora actividades que provocan la participación activa del alumnado en la valoración del aprendizaje.	
1. Sin participación	El alumnado interviene en el sistema de evaluación solo en calidad de agente evaluado.
2. Participación puntual	De forma puntual se plantean actividades de evaluación participativa (autoevaluación y/o evaluación por pares) que agilizan los procesos de corrección y devolución de respuesta.
3. Evaluación participativa: implementación	Se realizan actividades para que el alumnado aplique los criterios de evaluación definidos a la valoración de su aprendizaje y/o el de sus iguales.
4. Evaluación participativa: diseño e implementación	El alumnado interviene en la definición de las condiciones de la evaluación y en la aplicación de las mismas a la valoración de su aprendizaje y/o el de sus iguales.
5. Evaluación participativa: diseño, implementación y revisión	El alumnado interviene en el diseño y desarrollo del sistema evaluativo, así como en la revisión de sus principales ventajas y limitaciones.

Para acceder a las valoraciones de ambos sectores de la muestra se han configurado dos cuestionarios paralelos. Su validez de contenido reposa sobre el modelo teórico en el que se basan, elaborado a partir de referencias actuales y relevantes, entre las que se incluye la consulta experta (Bilbao, 2015). Los datos relativos a la validez de estructura interna proceden de la aplicación completa de las herramientas y han sido verificados por medio de un análisis factorial exploratorio por componentes principales. Las soluciones factoriales resultantes reducen las ocho dimensiones del modelo teórico original en seis factores, que explican el 77,43% de la varianza en el cuestionario docente y el 61,36% en el del alumnado.

Entendida como consistencia interna, el Alfa de Cronbach revela una fiabilidad satisfactoria para ambos instrumentos, con puntuaciones superiores al .85 en ambas sub-escalas (aceptación y cumplimiento).

Análisis

La información recopilada ha sido analizada por medio de estadísticos descriptivos y de inferencia bivariable. Por incumplimiento ostensible de los supuestos de normalidad, homocedasticidad y fortaleza de medida, se ha recurrido a técnicas estadísticas no paramétricas, ejecutadas por medio del programa SPSS (versión 23). Para contrastar la percepción de relevancia y cumplimiento de estudiantes y profesorado se ha aplicado la *U* de Mann Whitney, asumiendo un nivel de significación del 5% ($p < .05$). Para profundizar en las diferencias detectadas dentro del intervalo de confianza (IC = 95%), se han efectuado dos cálculos complementarios:

- Tamaño del Efecto (TE) por medio de la r de Rosenthal, que permite discriminar discrepancias de magnitud pequeña ($r \leq .1$), media ($.1 < r \leq .3$) o grande ($r \geq .5$) (Rosenthal, 1991).
- Prueba Chi-cuadrado y representación de tablas de contingencia, que permiten discriminar la concentración de estudiantes y profesorado en las diversas opciones de respuesta.

RESULTADOS

Aceptación de los retos evaluativos analizados

Analizados por medio de la *U* de Mann-Whitney, los promedios de importancia registrados no difieren significativamente ($p > .05$) entre estudiantes y profesorado. Las valoraciones producidas están muy próximas

entre sí y denotan una buena aceptación de los retos evaluativos planteados. Los promedios de respuesta oscilan en torno al 4, lo que denota que se consideran aspectos *Relevantes* (ver *Figura 2*).

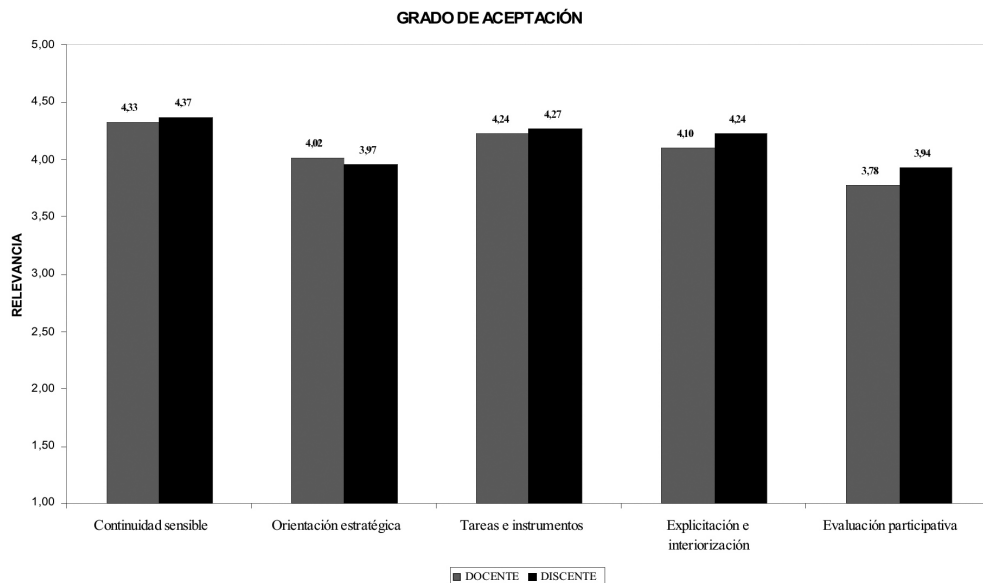


Figura 2. Importancia atribuida a los retos evaluativos por ambos sectores de la muestra

Los resultados representan el mismo patrón de respuesta en docentes y discentes. Para ambas fuentes la Continuidad sensible representa el indicador más importante, seguido de las Tares e Instrumentos puestos al servicio de la evaluación. En el extremo opuesto, la Evaluación participativa destaca como el menos relevante de los retos considerados. Salvo en la Orientación estratégica de la evaluación, los promedios de importancia más elevados se registran entre el alumnado.

Cumplimiento percibido en los retos evaluativos analizados

Las valoraciones registradas al calificar el grado de consecución percibido en los retos son menores y más dispares (*Figura 3*). Salvo en las Tareas e instrumentos para la evaluación de competencias, el alumnado manifiesta una visión más crítica, que sitúa el cumplimiento por debajo de la media teórica ($M < 3$). En el sector docente, eso solo sucede al describir la Participación discente en la evaluación.

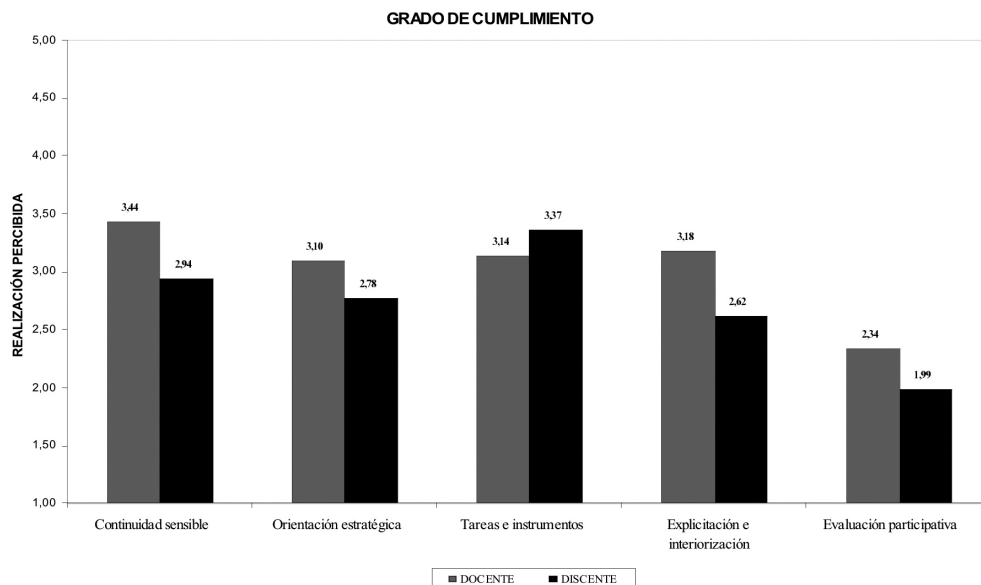


Figura 3. Cumplimiento atribuido a los indicadores por ambos sectores de la muestra

La *U* de Mann-Whitney revela diferencias estadísticamente significativas ($p > .05$) en tres de los indicadores considerados (ver Tabla 2), todas ellas sujetas a un tamaño de efecto pequeño.

La prueba Chi-cuadrado determina que docentes y discentes se sitúan en posiciones significativamente diferentes al describir el grado de Continuidad sensible ($p = .008$). En ambos casos, la alternativa más elegida es aquella que describe una práctica evaluativa muy adecuada (*continua y de intensidad graduada*), donde se concentra el 50% del profesorado participante y el 31,9% del alumnado (ver Tabla 3). Las diferencias principales se producen en las categorías inferiores, donde hay mayor proporción de estudiantes, y en la superior, donde la tendencia se invierte.

Tabla 2
Percepción del grado de consecución en función de la fuente

Estadísticos de contraste (U de Mann Whitney y Chi-cuadrado)						
Indicadores	U de Mann-Whitney	W de Wilcoxon	Z	Sig.	TE (r)	Chi-cuadrado (sig.)
Continuidad sensible	13166,500	260622,500	-3,082	,002	0,11 PEQ.	.008
Orientación estratégica	15223,500	263383,500	-1,640	,101	—	—
Tareas e instrumentos	15497,000	16772,000	-1,429	,153	—	—
Explicitación-interiorización	12211,000	259667,000	-3,747	,000	0,13 PEQ.	.001
Evaluación participativa	13826,500	257777,500	-2,591	,010	0,09 PEQ.	.067

Tabla 3
Percepción del grado de continuidad de la evaluación en función de la fuente: tabla de contingencia

Cumplimiento percibido		Docente	Discente	Total
Continuidad sensible	1. Final	Recuento 0	86	86
		% Fuente ,0%	12,2%	11,4%
	2. Final y proceso	Recuento 11	151	162
		% Fuente 22,0%	21,5%	21,5%
	3. Continua no atomizada	Recuento 10	212	222
		% Fuente 20,0%	30,2%	29,5%
	4. ...e intensidad graduada	Recuento 25	224	249
		% Fuente 50,0%	31,9%	33,1%
	5. ... y coordinada	Recuento 4	30	34
		% Fuente 8,0%	4,3%	4,5%
Total		Recuento 50	703	753
		% Fuente 100,0%	100,0%	100,0%

Nota: La segunda columna incorpora una versión simplificada de los descriptores de logro proporcionados para reflejar un grado de consecución inadecuado, poco adecuado, adecuado, muy adecuado y excelente.

La prueba Chi-cuadrado revela percepciones significativamente diferentes entre ambas fuentes al describir el grado de cumplimiento asociado a la Explicitación e interiorización ($p = .001$) (Tabla 2). Desde la perspectiva docente la publicidad de la evaluación está adecuadamente garantizada (ver Tabla 4). Señalada por el 42% del profesorado, la *explicitación escrita, junto a una presentación inicial y acciones de clarificación durante el curso* representa la opción más elegida; otro 26% declara proporcionar además *recursos de apoyo* para facilitar al alumnado la apropiación de los logros requeridos. En cambio, el 15,5% de las/os estudiantes acusa un cumplimiento inadecuado del reto, afirmando disponer solo de la información *explicitada* en el programa y/o guía docente. La alternativa más señalada es aquella que agrega una *presentación inicial* del sistema de evaluación, donde se sitúa el 36,8% de la respuesta discente. El nivel de realización *adecuado*, caracterizado por la realización de *acciones específicas de clarificación*, se suscribe en menor medida (25%).

Tabla 4
 Percepción del grado de explicitación y clarificación de la evaluación en función de la fuente: tabla de contingencia

	Cumplimiento percibido		Docente	Discente	Total
Explicitación/ clarificación	1. Explicitación	Recuento	0	109	109
		% Fuente	,0%	15,5%	14,5%
	2. Explicitación y presentación	Recuento	12	259	271
		% Fuente	24,0%	36,8%	36,0%
	3. Explicitación, presentación y clarificación en el curso	Recuento	21	176	197
		% Fuente	42,0%	25,0%	26,2%
	4. Clarificación apoyada	Recuento	13	110	123
		% Fuente	26,0%	15,6%	16,3%
	5. Actividades de interiorización	Recuento	4	49	53
		% Fuente	8,0%	7,0%	7,0%
Total		Recuento	50	703	753
		% Fuente	100,0%	100,0%	100,0%

Nota: La columna relativa al cumplimiento percibido incorpora una versión simplificada del contenido presentado en la rúbrica de evaluación para describir los cinco niveles de logro (inadecuado, poco adecuado, adecuado, muy adecuado y excelente) en este indicador en particular.

La Evaluación participativa representa el factor menos realizado para ambos sectores de la muestra. La prueba Chi-cuadrado no determina una

concentración significativamente diferente de docentes y discentes en las cinco opciones de respuesta ($p = .067$) (Tabla 2). De ello se deduce que las diferencias captadas por medio de la *U* de Mann-Whitney afectan más a la intensidad de la apreciación, que a su dirección. Para el profesorado la consecución de este indicador no alcanza el nivel de lo *adecuado* ($M = 2,34$; $s = 1,00$). El 22% reconoce *no realizar ninguna evaluación participativa*, el 36% emplearla *puntualmente* y el 30% de *forma limitada a lo largo del proceso*. El promedio de respuesta discente es menor aún ($M = 1,99$; $s = 1,06$), debido a una mayor concentración de casos en la categoría de *evaluación no participativa* (cumplimiento *inadecuado*, 44%). El porcentaje de estudiantes que reconoce una *participación puntual* o una *intervención limitada durante el proceso* (corrección de tareas fundamentalmente) desciende al 26,8% y al 20,3%, respectivamente.

DISCUSIÓN DE RESULTADOS

Las visiones puestas de manifiesto por estudiantes y profesorado toman distancia de las disfuncionalidades evaluativas clásicas, pero coinciden en determinar un evidente desajuste entre la importancia y el cumplimiento atribuido a los indicadores analizados. Su respuesta difiere de acuerdo a una tendencia, ampliamente reportada en estudios anteriores (Gutiérrez et al., 2011; Martínez-Mínguez y Flores, 2014; Tejedor, 1999; Romero, Castejón, y López, 2015), donde el alumnado ofrece mayores estimaciones sobre la importancia y puntuaciones menores al juzgar el grado de realización.

Dentro del patrón de respuesta descrito, las diferencias estadísticamente significativas detectadas son de tamaño pequeño y afectan únicamente al cumplimiento atribuido por ambas fuentes a los indicadores Evaluación participativa, Explicitación y clarificación y Continuidad sensible.

Para estudiantes y profesorado, la Evaluación participativa constituye el indicador menos relevante y menos llevado a cabo, configurando un área de mejora evidente, pero no prioritaria. A pesar de la vinculación natural de los Grados de Educación Infantil y Primaria con la capacitación evaluativa, predomina la heteroevaluación y ambas fuentes describen una intervención discente puntual, centrada principalmente en tareas de corrección durante el proceso formativo. Esta incidencia limitada de la evaluación participativa ya ha sido referida en otros estudios desarrollados en el ámbito específico de la formación básica del profesorado a nivel nacional (Arribas et al., 2010; Gutiérrez et al., 2013) e internacional (Gallardo-Fuentes y Carter, 2016; Gallardo-Fuentes et al., 2017). Representa una conclusión recurrente, incluso cuando estudiantes y profesorado reconocen beneficios pedagógicos y un mejor rendimiento al emplear formatos de evaluación participativa (Ibarra y

Rodríguez, 2014; Romero et al., 2015). De acuerdo con las virtudes descritas en la literatura específica, esta tendencia conlleva el desaprovechamiento de una vía privilegiada para contribuir al desarrollo de la autonomía y la capacidad evaluativa del alumnado (Álvarez, 2008; Brown, 2015).

Confirmando la discrepancia constatada en estudios realizados en el mismo contexto formativo (Arribas et al., 2010; Gutiérrez et al., 2011; Martínez et al., 2012), el profesorado reconoce un desarrollo práctico significativamente mayor al juzgar el grado de Explicitación e interiorización de los criterios de evaluación. De acuerdo con la mejora de la transparencia evaluativa que acusan Padilla et al. (2010) e Ibarra y Rodríguez (2010), docentes y discentes admiten la satisfacción de un primer nivel informativo, caracterizado por la explicitación escrita de los criterios y condiciones de la evaluación al comienzo del curso. Las discrepancias se producen al cuestionar el desarrollo de acciones específicas para profundizar en su comprensión. Mientras que el profesorado afirma llevar a cabo actividades específicas para facilitar el recuerdo, comprensión y profundización de los criterios establecidos, las/os estudiantes describen un cumplimiento mínimo que va poco más allá de la explicitación inicial contemplada en los referentes normativos. Sin embargo, más que información, la literatura específica insiste en la necesidad de ofrecer al alumnado apoyos que le capaciten para entender el significado de los objetivos y evaluar su situación de partida y su progreso respecto a los mismos. Para ello caben alternativas diversas y complementarias. Junto al *feedback* y el empleo de rúbricas (Álvarez y Olivera-Smith, 2014; Cabanal y Margalef, 2017), las fórmulas de evaluación participativa mencionadas en el indicador anterior representan una vía preferente para reducir las dificultades de descodificación de los estándares de logro propios de la disciplina (Price, Handley, Miller, & O'Donovan, 2010; Sambell, 2013).

De acuerdo con Pérez et al. (2013) y Hortigüela et al. (2015b) la Continuidad de las demandas evaluativas representa un poderoso factor para favorecer la organización y regulación del aprendizaje. En la muestra estudiada destaca como el indicador más relevante para ambas fuentes, pero el nivel de cumplimiento constatado podría calificarse como literal, sin indicios que apunten a su verdadera explotación pedagógica. Así, estudiantes y profesorado reconocen realizar tareas evaluativas en las tres fases del proceso, aunque no se advierte una graduación deliberada ni coordinación de las mismas. De modo similar, existen estudios que constatan una sobrecarga final en sistemas de evaluación continua, particularmente desde el punto de vista discente (Hamodi y López, 2012; Romero et al., 2015). Estas conclusiones sugieren el predominio de una visión prioritariamente instrumental de la continuidad evaluativa, basada en el número de pruebas y su distribución en el tiempo. Por el momento poco representado en la investigación, permanece el reto de alcanzar una

concepción más pedagógica, que incida en el impacto formativo de la sucesión de tareas evaluativas en la construcción del aprendizaje.

A pesar del avance que indican los resultados resumidos, el refuerzo de la misión pedagógica de la evaluación sigue representando un área mejorable pendiente de resolución. Existen evidencias de que el uso de la evaluación formativa se asocia a mejores resultados de aprendizaje (Brown, 2015; Fraile et al., 2013; Gallardo-Fuente y Carter, 2016) y estudiantes y profesorado coinciden en admitir su relevancia (Bilbao, 2015; Romero et al., 2015). Sin embargo, los cambios en el grado de explicitación, la continuidad, los agentes o los instrumentos de evaluación siguen más dirigidos a la acreditación de resultados que a la estimulación didáctica del proceso de construcción del aprendizaje.

De esta situación se concluye que la consolidación de un nuevo paradigma evaluativo excede la articulación de cambios parciales y puntuales, debiendo afectar a las convicciones profundas sobre la evaluación (Crossouard, 2010). Tales concepciones influyen significativamente en la forma de orientar y ejercer la evaluación, de modo que toda transformación de la praxis evaluativa debe partir de su comprensión y transformación (Hidalgo y Murillo, 2017).

A pesar de la fecunda línea de trabajo que se está consolidando en torno a ellas, la detección y modificación de las concepciones evaluativas representa una empresa compleja por su naturaleza implícita y su resistencia al cambio (Cruz, Pozo, Huarte, y Scheuer, 2006). La capacitación evaluativa inicial y permanente representa la alternativa de actuación más reseñada en ese camino (Cano y Ion, 2012; Ibarra y Rodríguez, 2010; López y Palacios, 2012; Stiggins, 1999), aunque dentro de un marco determinado de condiciones. Más que a incrementar el repertorio de técnicas e instrumentos, las iniciativas formativas deberían procurar la construcción de una imagen alternativa de la evaluación, que modifique las representaciones profundas que explican la prevalencia de su impronta acreditativa (Knight, 2005; Monereo y Castelló, 2009). Se trataría, en definitiva, de provocar las experiencias oportunas, para que el profesorado haga de la evaluación un recurso para regular la profundidad y calidad del aprendizaje discente (Asún et al., 2017). Desde esta perspectiva, Christoforidu et al. (2014) aportan que el alumnado de las/os docentes con mayor habilidad evaluativa obtienen mejores resultados de aprendizaje.

La superación profunda de los retos de mejora que persisten en la evaluación resulta especialmente necesaria en la formación básica de maestras/os y profesorado (López y Palacios, 2012; Palacios y López, 2013). Aunque también son afectadas por el contexto y las presiones socio-políticas,

buena parte de las concepciones evaluativas derivan de los aprendizajes inconscientes sobre evaluación contruidos en la etapa discente (Cruz et al., 2006; Prieto y Contreras, 2008). En consecuencia, este espacio representa un enclave estratégico para contribuir a la superación de las deficiencias clásicas y generar entre el futuro profesorado nuevas prácticas evaluativas que repercutan en otros niveles del sistema educativo.

NOTAS

- 1 El modelo original incorpora 14 indicadores más para la delimitación de otras cuatro dimensiones evaluativas: Atención a la diversidad, Implementación, Análisis e interpretación de evidencias, Regulación de la enseñanza y el aprendizaje y Metaevaluación.

REFERENCIAS BIBLIOGRÁFICAS

- Álvarez, I. (2008). La coevaluación como alternativa para mejorar la calidad del aprendizaje de los estudiantes universitarios: Valoración de una experiencia. *Revista Interuniversitaria de Formación del Profesorado*, 63(22,3), 127-140.
- Álvarez, I., y Olivera-Smith, M. (2014). Autorregulados. Sistemas basados en guías y rúbricas que promueven la autorreflexión. En C. Monereo (Coord.), *Enseñando a enseñar en la universidad. Formación del profesorado basada en incidentes críticos* (pp. 188-216). Barcelona: Octaedro.
- Arribas, J., Carabias, D., y Monreal, I. (2010). La docencia universitaria en la formación inicial del profesorado. El caso de la Escuela de Magisterio de Segovia. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 13(3), 27-25. Recuperado de <https://goo.gl/d1yM9m>
- Asún, S., Rosario, M., y Chivite, M. (2017). Exploración de sistemas de evaluación formativa entre estudiantes universitarios en la provincia de Huesca. *Apunts. Educación física y deporte*, 127, 52-58. doi: [http://dx.doi.org/10.5672/apunts.2014-0983.es.\(2017/1\).127.05](http://dx.doi.org/10.5672/apunts.2014-0983.es.(2017/1).127.05)
- Bilbao, A. (2015). *Competencia evaluativa del profesorado universitario de los Grados de Educación Infantil y Primaria* [Tesis doctoral]. Universidad de Deusto, Bilbao.
- Bretones, A. (2008). Participación del alumnado de educación superior en su evaluación. *Revista de Educación*, 347, 181-202. Recuperado de <https://goo.gl/tVeNS7>
- Brew, C., Riley, P., & Walta, C. (2009). Education students and their teachers: Comparing views on participative assessment practices. *Assessment & Evaluation in Higher Education*, 34(6), 641-657. doi: 10.1080/02602930802468567
- Brown, S. (2015). La evaluación auténtica: el uso de la evaluación para ayudar a los estudiantes a aprender. *RELIEVE*, 21(2). doi: 10.7203/relieve.21.2.7674
- Buscà, F., Pintor, P., Martínez, L., y Peire, T. (2010). Sistemas y procedimientos de evaluación formativa en docencia universitaria: Resultados de 34 casos aplicados durante el curso académico 2008-2009. *Estudios Sobre Educación*, 18, 255-276. Recuperado de <https://goo.gl/dRD63N>
- Canabal, C., y Margalef, L. (2017). La retroalimentación: la clave para la evaluación orientada al aprendizaje. *Profesorado*, 2, 149-170.
- Cano, E., y Ion, G. (2012). Prácticas evaluadoras en las universidades catalanas: Hacia un modelo centrado en competencias. *Estudios Sobre Educación*, 22, 155-177. Recuperado de <https://goo.gl/eiDgYA>
- Christoforidou, M., Kyriakides, L., Antoniou, P., & Creemers, B. (2014). Searching for stages of teacher's skills in assessment. *Studies in Educational Evaluation*, 40, 1-11. doi:10.1016/j.stueduc.2013.11.00
- Cid-Sabucedo, A., Pérez-Abellás, A., y Zabalza, M. (2009). Las prácticas de enseñanza declaradas de los «mejores profesores» de la Universidad de Vigo. *Revista Electrónica de Investigación y Evaluación Educativa*, 15(2). Recuperado de <https://goo.gl/PDvCaG>

- Crossouard, B. (2010). Reforms to higher education assessment reporting: Opportunities and challenges. *Teaching in Higher Education*, 35(3), 247-258. doi: 10.1080/13562511003740809
- Cruz, M., Pozo, J., Huarte, M., y Scheuer, N. (2006). Concepciones de enseñanza y prácticas discursivas en la formación de futuros profesores. En J. Pozo et al. (Eds.), *Nuevas formas de pensar la enseñanza y el aprendizaje. Concepciones de profesores y alumnos* (pp. 359-371). Barcelona: Graó.
- Fernández, M. (1989). *Así enseña nuestra universidad. Hacia una construcción crítica de una didáctica universitaria*. Madrid: Universidad Complutense de Madrid.
- Flórez, J. (1999). *Evaluación de la calidad de la docencia*. León: Universidad de León.
- Fraile, A., López, M., Castejón, F., y Romero, R. (2013). La evaluación formativa en docencia universitaria y el rendimiento académico del alumnado. *Aula Abierta*, 41(2), 23-34.
- Gallardo-Fuentes, y Carter, B. (2016). La evaluación formativa y compartida durante el prácticum en la formación inicial del profesorado: análisis de un caso en Chile. *Retos. Nuevas tendencias en Educación Física Deportes y Recreación*, 29, 258-263. doi: 10.7203/relieve.21.1.5169
- Gallardo-Fuentes, F., López-Pastor, V., y Carter, B. (2017). ¿Hay evaluación formativa y compartida en la formación inicial del profesorado en Chile? Percepciones del alumnado, profesorado y egresados de una universidad. *Psychology, Society & Education*, 9(2), 227-238. doi: 10.25115/psyse.v9i2.699
- Gil-Flores, J. (2012). La evaluación del aprendizaje en la universidad según la experiencia de los estudiantes. *Estudios Sobre Educación*, 22, 133-153. Recuperado de <https://goo.gl/RtcHCt>
- Gutiérrez, C, Pérez, A., y Pérez, M. (2013). Percepciones de profesores, alumnos y egresados sobre los sistemas de evaluación en estudios universitarios de formación del profesorado de educación física. *Ágora para la Educación Física y el Deporte*, 15(2), 130-151. Recuperado de <https://goo.gl/d8WgKu>
- Gutiérrez, C., Pérez, A., Pérez, M., y Palacios, A. (2011). Percepciones de profesores y alumnos sobre la enseñanza, evaluación y desarrollo de competencias en estudios universitarios de formación de profesorado. *Cultura y Educación*, 23(4), 499-514. doi: 10.1174/113564011798392451
- Hamodi, C., Y López, T. (2012). La evaluación formativa desde la perspectiva del alumnado y egresados. *Psychology, Society, & Education*, 4(1), 103-116. Recuperado de <https://goo.gl/Jnc2k5>
- Hidalgo, N., y Murillo, F. (2017). Las concepciones sobre el proceso de evaluación del aprendizaje de los estudiantes. *REICE*, 15(1), 107-128. doi: 10.15366/reice2017.15.1.007
- Hortigüela, D., Pérez, A., y Abella, V. (2015a). Perspectiva del alumnado sobre la evaluación tradicional y la evaluación formativa. Contraste de grupos en las mismas asignaturas. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 13(1), 35-48. Recuperado de <https://goo.gl/4NGVrS>
- Hortigüela, D.; Pérez, A., y López, V. (2015b). Implicación y regulación del trabajo del alumnado en los sistemas

- de evaluación formativa en educación superior. *RELIEVE*, 21(1). doi: 10.7203/relieve.21.1.5171
- Ibarra, M.S., y Rodríguez, G. (2010). Aproximación al discurso dominante sobre la evaluación del aprendizaje en la universidad. *Revista de Educación*, 351, 385-407. Recuperado de <https://goo.gl/n7BaL9>
- Ibarra, M.S., y Rodríguez, G. (2014). Modalidades participativas de evaluación: un análisis de la percepción del profesorado y de los estudiantes. *Revista de Investigación Educativa*, (32), 2, 339-362.
- Knight, P. (2005). *El profesorado de educación superior. Formación para la excelencia*. Madrid: Narcea.
- López, V. (2004). Evaluación, calificación, credencialismo y formación inicial del profesorado: Efectos y patologías generadas en la enseñanza universitaria. *Revista Interuniversitaria de Formación del Profesorado*, 18(3), 221-232.
- López, V., y Palacios, A. (2012). Percepción de los futuros docentes sobre los sistemas de evaluación de su aprendizaje. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 13(3), 317-341. Recuperado de <https://goo.gl/Ck1Gnf>
- Martínez, L., Castejón, F., y Santos, M. (2012). Diferentes percepciones sobre la evaluación formativa entre profesorado y alumnado en la formación inicial de educación física. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 15(4), 57-67. Recuperado de <https://goo.gl/rqTb28>
- Martínez-Mínguez, L., y Flores, G. (2014). Profesorado y egresados ante los sistemas de evaluación del alumnado en la formación inicial en la formación inicial del maestro en educación infantil. *Revista Digital de Investigación en Docencia Universitaria*, 1, 32-50. Recuperado de <https://goo.gl/eNWrvA>
- Monereo, C., y Castelló, M. (2009). La evaluación como herramienta de cambio educativo: Evaluar las evaluaciones. En C. Monereo (Coord.), *Pisa como excusa. Repensar la evaluación para cambiar la enseñanza* (pp. 13-30). Barcelona: Graó.
- Nieto, S. (2000). El discurso del profesorado universitario sobre la evaluación del aprendizaje de los alumnos como estrategia de innovación y cambio profesional: Exposición y análisis de una experiencia. *Revista de Educación*, 322, 305-324.
- Olmos, S., y Rodríguez, M. (2010). Diseño del proceso de evaluación de los estudiantes universitarios españoles: ¿Responde a una evaluación por competencias en el Espacio Europeo de Educación Superior? *Revista Iberoamericana de Educación*, 53(1). Recuperado de <https://goo.gl/kyUFSP>
- Padilla, M., Gil, J., Rodríguez, J., Torres, J., y Clarés, J. (2010). Evaluando el sistema de evaluación del aprendizaje universitario: Análisis documental de la Universidad de Sevilla. *Revista Iberoamericana de Educación*, 53(3). Recuperado de <https://goo.gl/L6wDR3>
- Palacios, A., y López, V. (2013). Haz lo que digo pero no lo que hago: Sistemas de evaluación del alumnado en la formación inicial del profesorado. *Revista de Educación*, 361, 279-305. doi: 10.4438/1988-592X-RE-2011-361-143
- Pereira, D., Flores, M., & Niklasson, L. (2016). Assessment revisited: a review of research in Assessment an Evaluation in Higher Education.

- Assessment & Evaluation in Higher Education*, 41(7), 1008-1032. doi: 10.1080/02602938.2015.1055233
- Pérez, M., Reyes, M., Palma, M., y Rafel, E. (2000). La evaluación de la calidad del aprendizaje en la universidad. *Infancia y Aprendizaje*, 23(91), 5-30. doi: 10.1174/021037000760087847
- Price, M., Handley, K., Miller, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, 35(3), 277-289. doi: 10.1080/02602930903541007
- Prieto, M., y Contreras, G. (2008). Las concepciones que orientan las prácticas evaluativas: Un problema a desvelar. *Estudios Pedagógicos*, 34(2), 245-262.
- Rodríguez, G. (Coord.). (1998). La evaluación desde la perspectiva del alumnado universitario: Propuestas para la mejora. *Revista de Enseñanza Universitaria*, número extraordinario, 223-234.
- Rodríguez, G., Ibarra, M., Gallego, B., Gómez, M., y Quesada, V. (2012). La voz del estudiante en la evaluación del aprendizaje: Un camino por recorrer en la universidad. *RELIEVE*, 18(2). doi: 10.7203/relieve.18.2.1985
- Romero, R., Castejón, J., y López, V. (2015). Divergencias del alumnado y del profesorado universitario sobre las dificultades para aplicar la evaluación formativa. *RELIEVE*, 21(1). doi: 10.7203/relieve.21.1.5169
- Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Newbury Park, CA: Sage.
- Sabirón, F., y Arraiz, A. (2013). Aprendiendo de la evaluación: Decálogo para la evaluación auténtica de competencias profesionales a través del portafolio. *Revista Iberoamericana de Evaluación Educativa*, 6(1), 135-152. Recuperado de <https://goo.gl/ua3ZXF>
- Sambell, K. (2013). Involving students in the scholarship of assessment: Student voices on the feedback agenda for change. En S. Merry, M. Price, D. Carless, & M. Taras (Eds.), *Reconceptualising feedback in higher education. Developing dialogue with students* (pp. 80-91). London: Routledge.
- Sayós, R. (2010). Evaluación continuada. En A. Parcerisa (Coord.), *Ejes para la mejora docente en la universidad* (pp. 126-132). Barcelona. Octaedro-ICE UB.
- Stiggins, R. (1999). Barriers to effective student assessments. *Education Digest*, 64(6), 25-29.
- Tejedor, F. (dir.) (1999). *Las estrategias utilizadas por los profesores universitarios para la evaluación del aprendizaje de los alumnos: Análisis y valoración de la situación actual, propuesta y seguimiento de experiencias de cambio dirigidas a la mejora de la calidad de la enseñanza y la profesionalización docente*. Recuperado de <https://goo.gl/sLoCw2>
- Trillo, F., y Porto, M. (1999). La percepción de los estudiantes sobre su evaluación en la Universidad. Un estudio de caso en la Facultad de Ciencias de la Educación. *Innovación Educativa*, 9, 55-75. Recuperado de <https://goo.gl/HNTzbz>
- Villa, A., Campo, L., Arranz, S., Villa, O., y García, A. (2013). Valoración del profesorado de magisterio sobre el aprendizaje basado en competencias implantado. *Profesorado*, 17(3), 35-55. Recuperado de <https://goo.gl/Z3C3TL>

PERFIL ACADÉMICO Y PROFESIONAL DE LOS AUTORES

Ainhoa Bilbao Martínez. Licenciada en Pedagogía y doctora en Innovación educativa por la Universidad de Deusto. Es Profesora Titular Doctora del Departamento de Psicopedagogía en la Escuela Universitaria de Magisterio Begoñako Andra Mari. Su actividad docente se desarrolla en el ámbito de la Didáctica y la Organización educativa y sus líneas de investigación prioritarias son la Evaluación del aprendizaje, la Docencia universitaria y la Formación del profesorado.

Aurelio Villa Sánchez. Doctor en Pedagogía por la Universidad Pontificia de Salamanca. Es profesor emérito de la Facultad de Psicología y Educación de la Universidad de Deusto y Catedrático de Métodos de Investigación en Educación. Ocupó el cargo de Vicerrector de Innovación, Calidad e Innovación en dicha institución desde el 2000 al 2010. Sus numerosas publicaciones abordan prioritariamente el tema del liderazgo, la formación y evaluación del profesorado y la innovación educativa.

Dirección de los autores: Ainhoa Bilbao Martínez
C/ Barrainkua, 2 (Despacho 605)
48009 Bilbao
E-mail: abilbao@bam.edu.es

Aurelio Villa Sánchez
Avda. de las Universidades 24 (Despacho
219-D)
48007-Bilbao
E-mail: Aurelio.villa@deusto.es

Fecha recepción del artículo: 11. Octubre. 2017
Fecha modificación del artículo: 09. Mayo. 2018
Fecha aceptación del artículo: 17. Mayo. 2018
Fecha revisión para publicación: 20. Junio. 2018